STANDARDIZATION OF OTOLITH-BASED AGEING PROTOCOLS

FOR ATLANTIC BLUEFIN TUNA

David H. Secor[1], Robert Allman[2], Dheeraj Busawon[3], Benjamin Gahagan[4], Walter Golet[5], Elise Koob[5], Patricia Lastra Luque[6], and Matt Siskey[1]

SUMMARY

Experts engaged in otolith-based age determination for Atlantic bluefin tuna convened at a workshop to evaluate ageing errors associated with preparation and viewing protocols and interpretation of annuli. Images of serial sections of the same otoliths, viewed under different light (reflected v. transmitted) were interpreted by 6 experienced and 2 naïve readers. Results indicated biased error associated with the sectioning plane. Type of lighting did not contribute significant bias. Participants developed consensus interpretation criteria by examining and annotating annuli (opaque zones) on a Smart Board. On imaged otolith sections, a reference bar demarcating the range of otolith sizes associated with the first opaque zone aided interpretation of the first annulus. Annulus interpretations were most uncertain for the first 6 annuli. Interpretation of the last opaque zone, near the edge of the sectioned otolith, also varied substantially among readers. An ageing precision trial on the same set of bluefin tuna otoliths, contributed from three laboratories, indicated moderately high precision among the 8 experienced readers. Average ageing error was 5.8±2.8%. Further, size-at-age estimates conformed well to the Restrepo et al. 2010 growth model regardless of the level of reader experience. The group tasked itself in the coming year to (1) develop an annotated reader set of 200 images of otoliths that are representative of preparation quality, viewing approaches, and fish sizes; (2) measure annulus dimensions to improve interpretations for younger ages; (3) undertake additional precision trials in interpreting the most terminal opaque zone; and (4) conduct marginal increment analysis to further evaluate the timing of annulus formation and inform interpretations of opaque/translucent zones.

---

[1] Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, P.O. Box 38, Solomons, Maryland 20688 U.S. secor@umces.edu
[2] Panama City Laboratory, National Marine Fisheries Service
[3] Fisheries and Oceans Canada, St. Andrews Biological Station
[4] Massachusetts Division of Marine Resources
[5] University of Maine and Gulf of Maine Research Institute
[6] Spanish Institute of Oceanography

1. Introduction

Stock assessments for Atlantic bluefin tuna are age-structured, reliant on estimates of catch-at-age, age-based fishing selectivity, growth, and schedules of reproduction and mortality. Further, planned multi-stock assessment models will depend on age-specific movement estimates. Interpretations of annuli in fin spines and otoliths have provided estimates of ages for Atlantic bluefin tuna (Rodriguez-Marin et al. 2007) but until recently these interpretations have not been evaluated for accuracy, bias, and precision. Several laboratories in Europe, Canada, and US are now collecting otoliths for estimates of age and growth, population assignments, and movements. These and past collections are large and should provision future stock assessment parameters with thousands of age interpretations. In a timely effort to standardize otolith-based ageing, members of these laboratories convened in January 2013 to undertake age interpretation trials, arrive at consensus views on protocol and interpretation guidelines, and develop a reader set to promote quality control and assurance procedures between laboratories.

Error underlying the interpretation of annuli includes several classes (Campana 2001; Christman 2006). Model specification error relates to the validity of yearly annulus formation. Technical error addresses varying approaches to otolith preparation and viewing of annuli. Interpretation error addresses how individual readers vary in their interpretations of annuli. Assignment error is based on laboratory conventions in assigning age such as non-integer reporting (e.g., when a mean age is assigned for the same otolith), adjustments for month of capture, and precision thresholds. Error is also classified as (1) accuracy: closeness to the true value; (2) bias: a systemic error that causes a deviation from the expected mean; and (3) precision: the degree of repeatability. Longevity of large and old Atlantic bluefin tuna (20-30 years), estimated through interpretation of otolith annuli, was validated using bomb radiocarbon dating by Neilson and Campana (2008). Carbon dating, cage-holding studies and mark-recapture have verified aspects of yearly annulus formation in Southern bluefin tuna, which has a similar longevity and latitudinal distribution to Atlantic bluefin tuna (Kalish et al. 1996; Clear et al. 2000; Gunn et al. 2008). However, no study has yet verified the yearly formation of opaque and translucent zones in otoliths for fish more commonly occurring in Atlantic bluefin tuna fisheries (ages 3-20). This represents an important limitation as these younger ages are crucial to the accuracy of stock assessment parameters, and ageing interpretations are most variable for the first 2-6 annuli.

The current study addresses age bias and precision related to technical and interpretation errors. Atlantic bluefin tuna age and growth laboratories have developed differing approaches so we wished to test whether bias was due to the manner in which otoliths were sectioned (i.e., "V-plane" v. "Y-plane") and viewed (reflected v. transmitted light). We also evaluated precision for a representative set of otolith images. Although workshop trials could not yet test accuracy or assignment errors, consensus views were developed on ways to limit and evaluate these classes of error, which are presented in the Discussion.

2. Methods

Annuli in sectioned otoliths are interpreted and annotated by readers through identification of successive opaque zones, which are typically narrower and have contrasting optical properties (opaque under transmitted light) in comparison to adjacent translucent zones. Images supplied to readers were demarcated with a radial yardstick to aid in designating the first opaque zone but otherwise were not previously annotated (**Figures 1 and 2**). The radial yardstick was derived from mean measures of the first opaque zones observed in a sample of yearling otoliths (N=22).

Technical Error Trial: We used serial sections of 25 otoliths to expose annuli in one of two transverse section planes: a "V-plane" or a "Y-plane," which were variously preferred across laboratories (**Figure 1**). The latter plane had precedence in the literature (Hurley and Iles 1983) and provided a larger field over which to interpret annuli. The V-plane has precedence in Southern bluefin and bigeye tuna and contains the core corresponding to the earliest deposited material during the larval/juvenile period. Nested within this comparison were otoliths imaged under

reflected or transmitted light (**Figure 2**).   Samples comprised 103-283 cm curved fork length (CFL) Atlantic bluefin tuna captured in US North Carolina, Gulf of Maine, and Spanish trap fisheries.   The opaque zones of otolith annuli were interpreted by six experienced and two naïve readers. Experienced readers had >1 year experience in interpreting annuli in otoliths of Atlantic bluefin tuna or similar species. Naïve readers had no experience in interpreting annuli in Atlantic bluefin tuna or other species. For each reader, paired differences between Y and V planes were analyzed for bias (departure from zero). Bias due to reader and light type were evaluated through analysis of variance.

Interpretation Error Trial: A sample of 35 otoliths sectioned in both Y- and V-planes was interpreted once by each of 10 readers. Nested within the sample were three light types associated with three laboratory groups: transmitted, reflected, and reflected through a blue filter (**Figure 2**). Precision was estimated as Average Percent Error (APE) and Coefficient of Variation (CV) as follows,

$$APE_j = 100\% \times \frac{1}{R}\sum_{i=1}^{R}\frac{|x_{ij} - X_j|}{X_j}$$

and
$$CV_j = 100\% \times \frac{\sqrt{\sum_{i=1}^{R}\frac{(X_{ij}-X_j)^2}{R-1}}}{X_j}$$

where R=number of readers; x_{ij}=age estimate for reader i for otolith j; X_j=mean age for otolith j.

Contributions of laboratory group and reader-experience to precision were evaluated through analysis of variance. Error across assigned ages was evaluated through age bias plots, selecting ages from one of the most experienced readers as the reference for presumed age. To evaluate whether ages conformed to expected growth rate, size-at-age for each reader was plotted against the accepted SCRS Western Atlantic bluefin tuna growth model (Restrepo et al. 2010).

3. Results

Technical Error Trial: The choice of section plane introduced significant bias (N=200; p=0.001). The Y-plane resulted in an estimated ages that were 0.77 years higher than the V-plane sections (**Figure 1**). Experienced readers deviated slightly more (difference=0.90 years) than inexperienced readers (difference=0.38 years) but this difference was not significant (P=0.08). Interpretations of six of the eight readers showed levels of bias similar to the overall mean; the ages of two readers showed no bias between the plane types (**Figure 1**). In an analysis of variance, reader effects were significant (p=0.013) but the type light (reflected or transmitted) and its interaction with reader were nonsignificant (p>0.4).

Interpretation Error Trial: Average Percent Error was slightly higher for the entire group of 10 readers (APE=7.2%) than for the 8 experienced readers (APE=5.8%). For the experienced readers, the Coefficient of Variation was also lower than for the entire group (CV=7.8% v. 9.7%) (**Table 1; Figure 3**).   Precision varied between laboratories (p=0.03), but this difference was likely due to lower precision associated with ageing younger fish (**Figure 3**). An age biased plot detected no systemic bias in age estimates with increasing nominal age (**Figure 3**). Size-at-age plots for each reader showed strong overlap with the current growth curve for the Western Atlantic bluefin tuna stock (**Figure 4**).

4. Discussion

A chief objective of the workshop was to agree on preferred methods of preparation and observation to reduce error associated with age estimation across multiple laboratories. The significant bias between section planes for the same

otolith detected here was of sufficient amplitude that participants agreed that this aspect of preparation should be standardized. Because the Y-plane has precedence in the literature (Hurley and Iles 1983; Lee et al. 1983; Neilson and Campana 2008) this section type was selected by workshop participants and should be used for ongoing and future age determinations of Atlantic bluefin tuna. It is noteworthy that the current Restrepo et al. (2010) growth model relied upon ages derived from V-plane otolith sections. Any bias due to this section type would be expected to result in small to insignificant changes to this growth model, principally influencing the intercept t_0 because the bias was consistent across nominal age classes and was of relatively small amplitude (0.7 years). Further, the intercept function of the fitted growth model was most strongly influenced by juvenile length-frequency data in comparison to otolith ages, which principally influenced the fit to older (>3 years) juveniles and adults.

The type of light used in interpreting annuli was deemed largely a matter of preference and in workshop trials showed no detectable bias. Still, careful microscopy and image enhancement will improve precision and exchange of otolith images regardless of light type. Recommendations for improved viewing include (1) use of Photoshop © for image enhancement, annotation, and archiving; (2) use of large monitors (>70 cm diagonal) or Smart Boards in routine ageing and training; and (3) use of a first annulus "yard-stick" to assist in observing the first opaque zone.

In general, the ~6% ageing error between readers was low for a moderately long-lived species. Error was higher for fish with nominally younger ages, but this is in part due to APE and CV formulations which place mean age in the denominator and causes unavoidable inflation of calculated error for young fish. Age estimates did not systematically change with nominal age as one might expect if older or younger fish were more prone to interpretation bias. Further, size-at-age estimates conformed well to the Restrepo et al. (2010) growth model across all readers.

Participants with experience in quality control procedures in large age production laboratories indicated that a 10% APE was commonly implemented for species like Atlantic bluefin tuna (e.g., king mackerel) that are somewhat difficult to age. With improved standardization and experience, participants thought that a mean APE <5% was attainable, a level suggested by Campana (2001) as a common threshold level for production ageing laboratories. Chief sources of error identified by the group were (1) identification of the first 5-6 opaque zones; (2) identification of a final opaque zone near the edge of each section; and (3) seasonal interpretations of opaque and translucent zones. Tasks are now being undertaken by participants to address each of these issues.

A chief outcome of the workshop will be the development of an annotated reference set of a moderately large and representative sample of Atlantic bluefin tuna images. Such a reference set is essential for maintaining within and inter-laboratory consistency. Within laboratories, new readers will require training and evaluation before proceeding to routine ageing. Experienced readers will need to evaluate their internal consistency such as the tendency to drift or become increasingly rigid (canalized) in interpretations. Laboratories can employ quality assurance procedures by mixing sub-samples from a reference set into those samples under investigation. Between laboratory exchanges of a reference set will assure certain precision levels are achieved so as not to bias parameters and stock assessments, which often rely on age estimates provided from multiple laboratories. The "gold standard" for reference sets is a known-age sample (Campana 2001). Absent this, many groups have developed reference sets based on consensus ages by a group of experts: the current goal that we are pursuing.

A set of 100 otoliths from three laboratories has been selected stratified by fish size, and image quality. Each otolith section (Y-plane) has been image-enhanced under reflected and transmitted light (200 images total). We have designated one experienced reader as the nominal expert who will annotate and assign ages to all the images. These images will be then returned without annotation for each participant to independently annotate and assign ages. Analysis of precision and bias will follow and should precision be judged sufficient, the nominal expert's annotated images will be used as the reference set.

We have deferred decisions on conventions related to age assignment such as (1) should each otolith be interpreted by a single or multiple readers? (2) should fish be assigned fractional ages based upon multiple reads or depending on when it was captured in the season? and (3) should individual interpretations be excluded on the basis of precision thresholds? These issues are important in production ageing (e.g., GSMFC 2009). A further adumbrating issue is the fundamental question about the accuracy of age estimates for which additional dedicated research will be required.

5. Acknowledgements

6. Literature Cited

Campana, S. E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. Journal of Fish Biology 59(2):197-242.

Christman, M. 2006. Characterization and Incorporation of Uncertainty in Fisheries Management, p. 213-236. *In* Fisheries Ecosystem Planning for Chesapeake Bay. American Fisheries Society Monograph: Trends in Fisheries Science and Management 3, Bathesda.

Clear, N. P., J. S. Gunn, and A. J. Rees. 2000. Direct validation of annual increments in the otoliths of juvenile southern bluefin tuna, *Thunnus maccoyii*, by means of a large-scale mark-recapture experiment with strontium chloride. Fishery Bulletin 98:25-40.

Gulf States Marine Fisheries Commission. 2009. A Practical Handbook for Determining the Ages of Gulf of Mexico Fishes, Second Edition. Publication Number 167. Ocean Springs Mississippi. 157 p.

Gunn, J. S., and six coauthors. 2008. Age and growth in southern bluefin tuna, *Thunnus maccoyii* (Castelnau): Direct estimation from otoliths, scales and vertebrae. Fisheries Research 92(2-3):207-220.

Hurley, P.C.F. and T.D Iles. 1983. Age and growth estimation of Atlantic bluefin tuna, *Thunnus thynnus*, using otoliths. Proc. Int. Workshop on Age Determ. of Ocean Pel. Fishes 1982:71-75.

Kalish, J. M., J. M. Johnston, J. S. Gunn, and N. P. Clear. 1996. Use of the bomb radiocarbon chronometer to determine age of southern bluefin tuna *Thunnus maccoyii*. Marine Ecology Progress Series 143:1-8.

Lee, D.W., E.D. Prince, and ME. Crow. Interpretations of growth bands on vertebrae and otoliths of Atlantic bluefin tuna *Thunnus thynnus.* Proc. Int. Workshop on Age Determ. of Ocean Pel. Fishes 1982:61-70.

Neilson, J. D., and S. E. Campana. 2008. A validated description of age and growth of western Atlantic bluefin tuna (*Thunnus thynnus*). Canadian Journal of Fisheries and Aquatic Sciences 65(8):1523-1527.

Radriguez-Marin, E. and nine others. 2007. Report of the 2006 ICCAT Workshop for Bluefin Tuna Direct Ageing. SCRS/2006/157. Col. Vol. Sci. Pap. ICCAT, 60 (4): 1349-1392.

Restrepo, V.R., G.A. Diaz, J.F. Walter, J. Neilson, S.E. Campana, D. Secor, and R.L. Wingate. 2010. An updated estimate of the growth curve of Western bluefin tuna. Aquatic Living Resources 23: 235-342.

Table 1. Precision estimates for a set of Atlantic bluefin tuna otoliths that was interpreted by 8 experienced and 2 inexperienced readers.

BFT – West Atlantic

	Range	Mean	N
Ageing Percent Error – all readers	0-12.5	7.2	35
Ageing Percent Error – experienced readers	2.3-21.2	5.8	35
Coefficient of Variation – all readers	2.7-37.2	9.7	35
Coefficient of Variation – experienced readers	0-16.9	7.8	35

Figure 1. Comparisons of serial sections of same otolith, representing a V-plane (A) and a Y-plane (B). The V-plane is taken just anterior to the tip of the anti-rostrum and the Y section is taken posteriorly so that it includes a portion of the anti-rostrum (the base of the Y). Otolith images are annotated with filled blue circles indicating opaque zones of each annulus. Box whisker plots of paired differences between counts of annuli in Y- versus V-planes (C) indicated that the Y-plane shows a consistent positive bias relative to the V plane. A first annulus "yardstick" showed the range of distances associated with the first opaque zone.
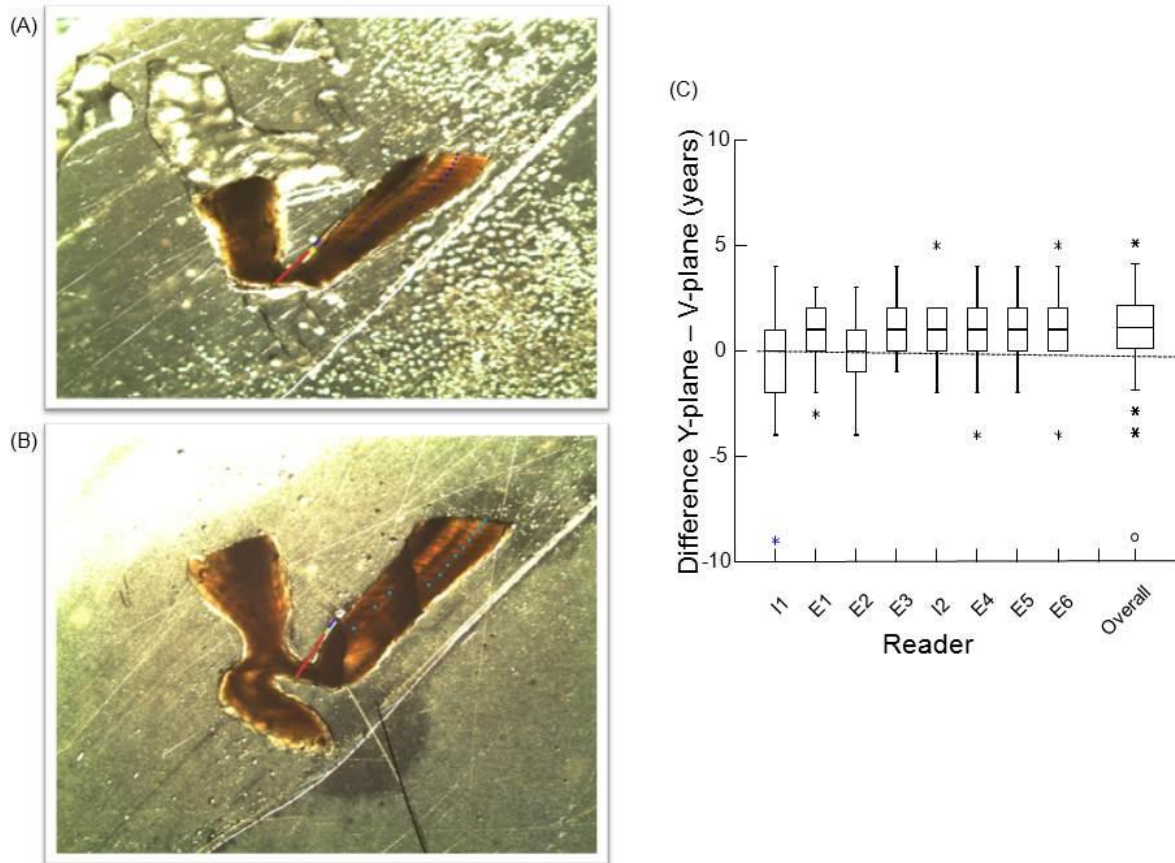
Figure 2. Comparisons of preferred light types between three laboratories for (A) transmitted light, (B) reflected white light, and (C) reflected light through a blue filter (aka purple haze). Otolith images are annotated with filled circles indicating opaque zones of each annulus. A first annulus "yardstick" shows the range of distances associated with the first opaque zone to guide its placement.
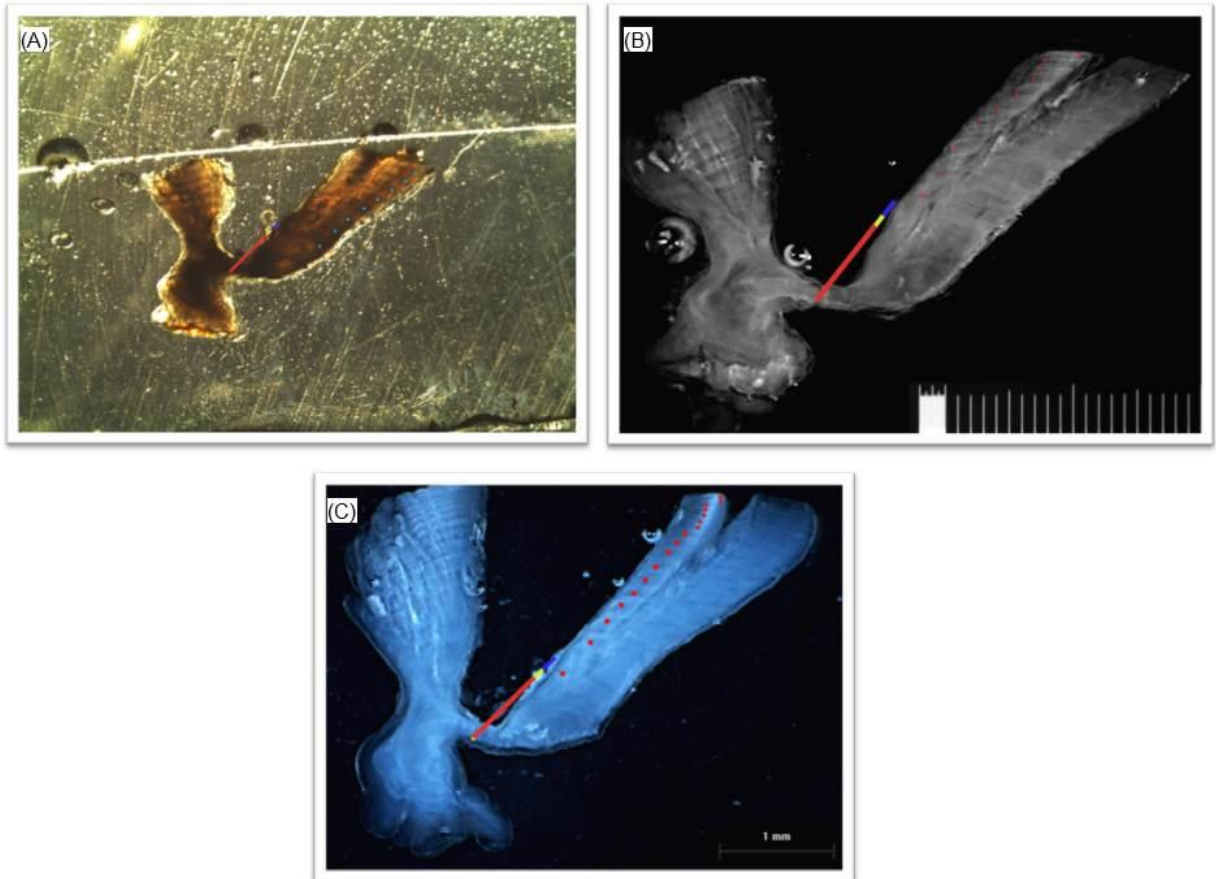
Figure 3. Precision of age estimates between readers for the same set of 35 otoliths. (A) distribution of Average Percent Error for eight experienced readers; (B) Average Percent Error for experienced readers plotted by laboratory against the nominal reference age (determined from a single experienced reader); (C) Coefficient of Variation versus Average Percent Error by plotted by laboratory; (D) Age bias plot of nine experienced and naïve readers against the nominal reference age.
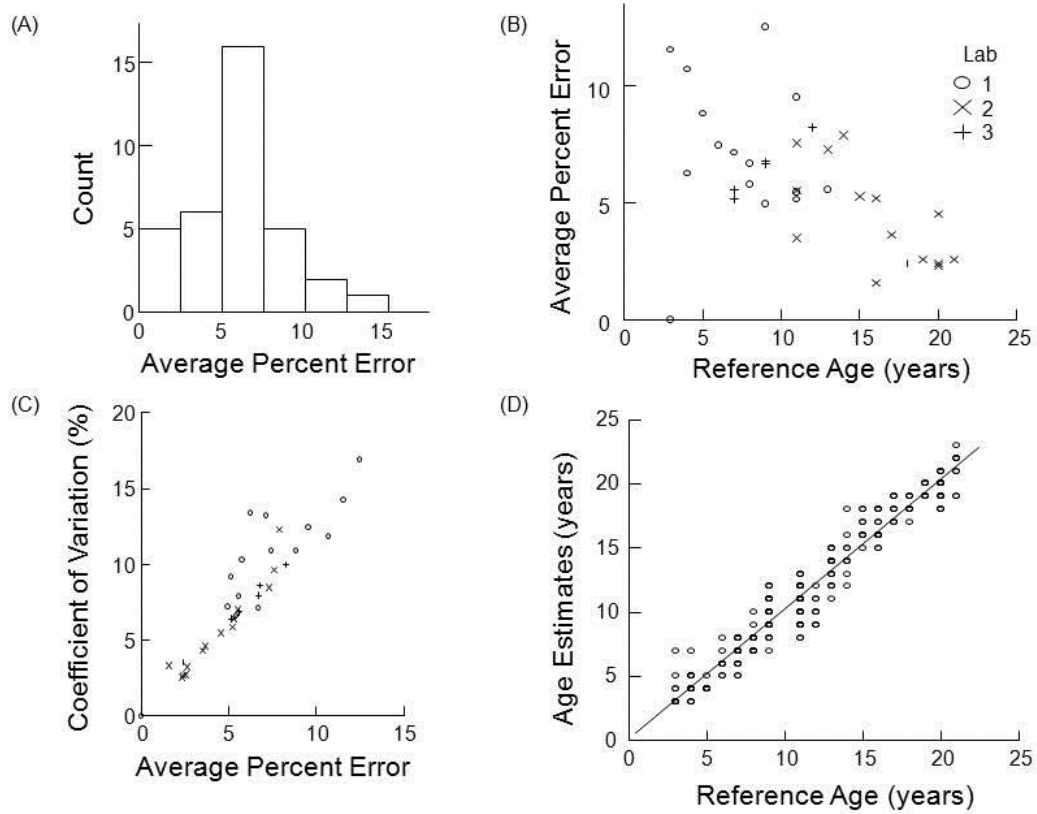
Figure 4. Estimated size-at-age by reader. Each plot shows estimated ages as red triangles (N=35) overlaying the Restrepo et al. (2010) growth curve (black squares). CFL=curved fork length.